

Underwater Optical Measurement of Position and
Orientation Using a Prepared, Passive Target

by

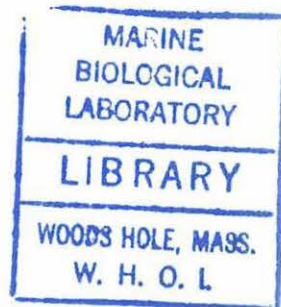
Niel Raymond Nickolaisen

B.S., Physics, Utah State University (1983)

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OCEAN ENGINEERING
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
and the
WOODS HOLE OCEANOGRAPHIC INSTITUTION

September, 1989



Copyright Niel Raymond Nickolaisen, 1989 All rights reserved.
The author hereby grants to MIT and WHOI permission to
reproduce and to distribute copies of this thesis
document in whole, of in part.

Signature of Author
Joint Program in Oceanographic Engineering
Massachusetts Institute of Technology
Woods Hole Oceanographic Institution
September 15, 1989

Certified by
Dr. Dana R. Yoerger
Woods Hole Oceanographic Institution
Thesis Supervisor

Certified by
Dr. Arthur B. Baggeroer
Massachusetts Institute of Technology

Accepted by
Dr. W. Kendall Melville
Chairman, Joint Committee for Oceanographic Engineering
Massachusetts Institute of Technology/
Woods Hole Oceanographic Institution



Underwater Optical Measurement of Position and
Orientation Using a Prepared, Passive Target

by

Niel Raymond Nickolaisen

Submitted to the Massachusetts Institute of
Technology/Woods Hole Oceanographic Institution Joint Program
in Oceanographic Engineering on September 15, 1989 in partial
fulfillment of the requirements for the degree of Master of
Science in Ocean Engineering.

Abstract

In order to provide an accurate and efficient method of docking, station keeping, and navigation for the JASON remotely operated vehicle, an optical tracking method is investigated. The method needs to be accurate enough for underwater control and fast enough to enable timely control. The method presented solves the camera location problem in a closed form and is accurate for accurate measurements of image plane coordinates. Target design criteria are discussed and a prepared, passive target is selected. Testing and error analysis reveal that the approach gives good results for camera lenses with focal lengths greater than 4.8mm. The effects of underwater use are discussed and a non-numerical method of compensating for the underwater effects is presented. The compensation method shifts the image plane coordinates toward the image plane center and gives results that are within acceptable error margins.

Thesis Supervisor: Dr. Dana R. Yoerger
Woods Hole Oceanographic Institution

Acknowledgements

I am very appreciative to Dr. Dana Yoerger for his help and patience. He provided the right insights at the right times. I would also like to thank Rikki Razdan and Alan M. Kielar at ISCAN and I hope that others may see the potential their tracking systems have in underwater applications. I thank the United States Navy, the Massachusetts Institute of Technology, and the Woods Hole Oceanographic Institution for the opportunity to participate in the Joint Program. Most of all, I thank my wife, Karen, for her encouragement and support.

Table of Contents

Chapter 1	Introduction	6
Chapter 2	Use of Perspective Transformations	
2.1	Detailed Analysis of the Use of Perspective Transformations to Determine ROV Position	10
2.2	Determining the Transformation Matrix	26
Chapter 3	Use of Prepared, Passive Targets	
3.1	Target Design Criteria	29
3.2	Target Spot Placement	32
Chapter 4	Description of the ISCAN Video Tracking System	34
Chapter 5	Testing	
5.1	Test Design	35
5.2	Camera Calibration	36
5.3	Test Results	37
Chapter 6	Underwater Effects	39
Chapter 7	Conclusions	44
Figures		45
Appendices		60
Bibliography		73

List of Figures

Figure 1	ARGO/JASON Navigation Referenced to Ship	45
Figure 2	JASON Navigation Referenced to ARGO	46
Figure 3	JASON Navigation Referenced to Seafloor	47
Figure 4	Displacement of Camera Coordinate System	48
Figure 5	Pan Angle Rotation	49
Figure 6	Tilt Angle Rotation	50
Figure 7	Swing Angle Rotation	51
Figure 8	Possible Target Spot Location	52
Figure 9	Test Target Base Plate	53
Figure 10	Error Versus Range (8mm lens in air)	54
Figure 11	Error Versus Range (uncorrected 8mm lens in water)	55
Figure 12	Error Versus Range (corrected 8mm lens in water)	56
Figure 13	Error Versus Range (12mm lens in air)	57
Figure 14	Error Versus Range (uncorrected 12mm lens in water)	58
Figure 15	Error Versus Range (corrected 12mm lens in water)	59

Chapter 1 Introduction

This paper and the associated research involve work done on a visual tracking and control system for the JASON remotely operated vehicle (ROV). JASON is part of the JASON/ARGO project. The problem, stated simply, is the following: Is it possible to visually track JASON and use information about its position to control JASON in station keeping, docking, and navigation procedures?

The key to accomplishing this task is being able to accurately determine the position of the ROV in its underwater environment. I attempt to find JASON's position by using a method of perspective transformations and by using a spot tracker invented by engineers at ISCAN Corp in Cambridge, MA.

The ARGO/JASON system is an advanced scientific tool that is designed to explore the ocean floor. The ARGO sled carries advanced sonar and video systems that scan the ocean floor. When the ARGO "discovers" something of interest and the system operators want to take a closer look at the "discovery", the JASON ROV is dispatched from a port in ARGO. Aboard ship, the operator maneuvers JASON manually or automatically using various acoustic navigation references.

Currently, JASON's position can be determined by several

different acoustic references. These are summarized in figures 1, 2, and 3.

The first is a low frequency long baseline system that is also used to navigate ARGO and the surface ship. While the system works over a long range (up to 10,000 m), it suffers from modest precision (plus or minus 10 m) and a low update rate (one fix every ten seconds). Also, JASON can easily be out of sight of the transponders when operating in rough terrain.

The second navigation reference is a high frequency system called SHARPS. Like the low frequency system, it uses the travel time of acoustic pulses to determine position. Due to the high bandwidth of SHARPS, range precision is about 1 cm. The high frequency of SHARPS (300 kHz) permits update rates of up to 10 fixes per second. A short baseline on JASON is used to position JASON relative to ARGO. SHARPS transponders, currently in development, will be used to navigate JASON relative to a fixed net on the seafloor.

It would simplify the use and control of JASON if it were possible to use JASON's on board cameras to determine JASON's position relative to some fixed target or landmark. This position information (and information about the time rate of change of position) could then be used to automatically

control JASON's thrusters to maneuver it into or keep it in a desired position. This would enable the operator to use the manipulator arm without having to also maintain JASON's position against any currents or disturbances. Knowing JASON's position relative to some target or landmark would provide a method of passive navigation. Locating JASON visually relative to an object would be a complementary alternative to the acoustic systems. While inappropriate for mapping and survey, visual navigation could allow JASON to hover close to objects without installing a seafloor net. This capability could be used to aid in manipulation or close-up inspection.

Developing a visual method of controlling JASON depends on whether it is possible to accurately and quickly determine JASON's position. Stated differently, if we know the shape or characteristics of a landmark or target and how that landmark or target looks to JASON's cameras, can we accurately determine how far away and at what angles JASON is from the landmark or target? Furthermore, can we obtain this information quickly enough to control JASON?

These questions were investigated in several steps. First, study the methods of finding a camera's position relative to a target and choose the method that can

solve the problem quickly and accurately enough to meet the control needs of JASON. Second, based on the method chosen above, design a target or teach the system a landmark that can be used as a reference. Third, test and implement the method and system.

Based on a review of the literature [Duda, 1973], [Ganapathy, 1984], [Haralick, 1980], it appears that perspective transformations provide the best mathematical approach to determining ROV position. The method described here looks at the relationship between target spot coordinates and those spots as they appear on the camera image plane and then uses this relationship to get a closed form solution to the problem. This solution method leads to the selection of a target composed of four spots on a plane.

Chapter 2

2.1 Detailed Analysis of the Use of Perspective Transformations to Determine ROV Position

To visually control the ROV, one must know its position relative to the target. This is found by decomposing the problem into a series of coordinate transformations and their associated transformation matrices. The transformations relate the position and orientation of the camera (and ROV) to some fixed point (the target). These transformations move the coordinate axes from the target to the camera center and rotate the axes to line up along the camera's line of sight.

Two coordinate systems will be used, one for the target and one for the camera (ROV). Let the target coordinate system be labeled X,Y,Z . Let the camera be centered at a point S with a coordinate system labeled X',Y',Z' . Let the point P be where the line of sight of the camera intercepts the X,Y plane. Let the point O be where the line of sight of the camera intercepts the image plane (the U,V plane). Point O is the principal point. Let the line $S'P$ represent the projection of SP onto the X,Y plane. These coordinate systems, points, and lines are shown in figure 4.

The first step of the coordinate transformation is performed by moving the camera coordinate axes to S . Next,

rotate the X', Y', Z' system around the Z' axis by an angle (l) (the pan angle) such that the Y' axis is parallel to $S'P$ (see figure 5). Now, rotate the X', Y', Z' system around the X' axis by an angle (m) (the tilt angle) such that the Z' axis lines up with SP (see figure 6). As a final step, rotate the X', Y', Z' system around the Z' axis by an angle (n) (the swing angle) such that the X' axis is parallel to the U axis of the image plane (see figure 4). After these transformations have been done, the X' axis will be parallel to the U axis of the image plane, the Y' axis will be either parallel or anti-parallel to the V axis of the image plane, and the Z' axis will be aligned with SP . Separately, there four steps can be summarized as follows:

1. Moving the origin to S :

$$[D] = \text{Displacement Matrix} = \begin{bmatrix} 1 & 0 & 0 & -X_s \\ 0 & 1 & 0 & -Y_s \\ 0 & 0 & 1 & -Z_s \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Where X_s, Y_s, Z_s are the coordinates (in the X, Y, Z system) of the point S .

2. Rotating X', Y', Z' about Z' (pan angle of l):

$$[R1] = \text{First Rotation Matrix} = \begin{bmatrix} \cos(l) & \sin(l) & 0 & 0 \\ -\sin(l) & \cos(l) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

3. Rotating X', Y', Z' about X' (tilt angle of m):

$$[R2] = \text{Second Rotation Matrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \cos(m) & \sin(m) & 0 \\ 0 & -\sin(m) & \cos(m) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

4. Rotating X', Y', Z' about Z' (swing angle of n):

$$[R3] = \text{Third Rotation Matrix} = \begin{bmatrix} \cos(n) & 0 & -\sin(n) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(n) & 0 & \cos(n) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Combining the rotation steps gives a matrix that relates the three rotation angles:

$$[R] = [R3][R2][R1]$$

The elements of $[R]$ can be written as:

$$[R] = \begin{bmatrix} a & b & c & 0 \\ d & e & f & 0 \\ g & h & i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\text{where: } a = \cos(n)\cos(l) - \sin(l)\sin(m)\sin(n)$$

$$b = \cos(n)\sin(l) + \cos(l)\sin(m)\sin(n)$$

$$c = -\cos(m)\sin(n)$$

$$d = -\sin(l)\cos(m)$$

$$e = \cos(l)\cos(m)$$

$$f = \sin(m)$$

$$g = \cos(l)\sin(n) + \sin(l)\sin(m)\cos(n)$$

$$h = \sin(m)\sin(n) - \cos(l)\sin(m)\cos(n)$$

$$i = \cos(m)\cos(n)$$

The rotation matrix, $[R]$, can be combined with the displacement matrix, $[D]$, to obtain a matrix whose elements contain information about the position of the camera relative to the target. The resulting matrix will be called $[EXT]$ because it contains information extrinsic to the

camera.

$$[EXT] = [R][D]$$

$$[EXT] = \begin{bmatrix} a & b & c & p \\ d & e & f & q \\ g & h & i & r \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where a, b, c, d, e, f, g, h, and i are the same as before and:

$$p = -aX_s - bY_s - cZ_s$$

$$q = -dX_s - eY_s - fZ_s$$

$$r = -gX_s - hY_s - iZ_s$$

It is worth noting here that homogeneous coordinates are being used which is why there are four rows in the above matrices. Conversion back to the three dimensional coordinate system is done by dividing all other rows by the fourth (this will be covered in detail when discussing the closed form solution to the problem).

Because the measurements of camera position will use images projected onto the image plane, the camera center coordinates (X', Y', Z') must be converted to coordinates

centered on the image plane (U,V). Let u' and v' denote coordinates of a point on the image plane. u' and v' are related to the X',Y',Z' coordinate system by:

$$u' = (x' * F) / z'$$

$$v' = (y' * F) / z'$$

where F is the focal length of the camera and x',y',z' are the coordinates of the point on the image plane in the X',Y',Z' coordinate system.

Because measurements on the image plane will be rasterized (in pixels), it is necessary to scale the measurements to get them in terms of units that will be used to measure the position of the camera relative to the target. This can be done by including a factor of k_u (in the U direction) and a factor of k_v (in the V direction). k_u will be assumed positive but k_v can be either positive or negative.

Now, with the coordinates of the principal point, O , being labeled (u_o, v_o) , the relation between the X',Y',Z' coordinate system and the image plane coordinates, u and v , can be written as:

$$u = u_o + (k_u * x' * F) / z'$$

$$v = v_o + (k_v * y' * F) / z'$$

This expression can be simplified by letting $k_u * F$ be k_1 and $k_v * F$ be k_2 . Writing this in terms of a homogeneous coordinate system yields the result:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 & 0 & u_0 & 0 \\ 0 & k_2 & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix}$$

or,

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 & 0 & u_0 \\ 0 & k_2 & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

The matrix containing k_1 , k_2 , u_0 , v_0 will be called [INT] since it contains information intrinsic to the camera i.e. focal length, rasterization scaling, and principal point coordinates.

Previously, the X', Y', Z' coordinate system was related to the x, y, z , system by the [EXT] matrix:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = [\text{EXT}] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Also, the U,V coordinate system was related to the X',Y',Z' system with the [INT] matrix:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = [\text{INT}] \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

These matrices can be combined to relate the U,V coordinate system to the X,Y,Z coordinate system (where the target is located):

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = [\text{INT}] [\text{EXT}] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = [T] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

where $[T]$ is a 3×4 matrix whose elements are:

$$t_{11} = k_1 a + u_o g$$

$$t_{12} = k_1 b + u_o h$$

$$t_{13} = k_1 c + u_o i$$

$$t_{14} = k_1 p + u_o r$$

$$t_{21} = k_2 d + v_o g$$

$$t_{22} = k_2 e + v_o h$$

$$t_{23} = k_2 f + v_o i$$

$$t_{24} = k_2 q + v_o r$$

$$t_{31} = g$$

$$t_{32} = h$$

$$t_{33} = i$$

$$t_{34} = r$$

So, given the coordinates of a target spot and measuring the coordinates of this spot as projected on the image plane, the elements of $[T]$ that connect the coordinates can be determined. Knowing the elements of $[T]$, the position and orientation of the camera relative to target (X_s , Y_s , Z_s and

1, m, n) can be found.

This problem can be somewhat simplified when some information is known about the camera parameters (k_1 and k_2 or u_0 and v_0). If four target spots are used and the spots are placed in general position on a plane (let that plane be the plane where $z = 0$) then the general approach can be written as:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 0 \\ 1 \end{bmatrix}$$

where $i = 1, 2, 3, 4$

This leads to:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{14} \\ t_{22} & t_{22} & t_{24} \\ t_{31} & t_{32} & t_{34} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

Taking advantage of the use of a homogeneous coordinate system (dividing each element by the last element, r) the above matrix elements can be written as:

$$t_{11} = (k_1 a + u_o g) / r$$

$$t_{12} = (k_1 b + u_o h) / r$$

$$t_{14} = (k_1 p + u_o r) / r$$

$$t_{21} = (k_2 d + v_o g) / r$$

$$t_{22} = (k_2 e + v_o h) / r$$

$$t_{24} = (k_2 q + v_o r) / r$$

$$t_{31} = g / r$$

$$t_{32} = h / r$$

$$t_{34} = 1$$

Knowing x_i and y_i , and measuring u_i and v_i , the elements of $[T]$ can be found. Then, these values, k_1 and k_2 , and the system constraint equations can be used to determine the values of a , b , c , d , e , f , g , h , and i . p , q , and r can then be found. Once this is done, X_s , Y_s , Z_s , l , m , and n can be computed.

Because the rotation matrix, $[R]$, is a pure orthonormal matrix, $[R]^{-1}$ is the same as $[R]^T$ and the determinant of $[R]$ is equal to 1 [Kanattani, 1987]. These conditions give the following constraints on the elements of $[R]$:

$$a^2 + b^2 + c^2 = 1$$

$$d^2 + e^2 + f^2 = 1$$

$$g^2 + h^2 + i^2 = 1$$

$$a = ei - hf$$

$$b = fg - di$$

$$c = dh - eg$$

$$d = hc - bi$$

$$e = ai - gc$$

$$f = bg - ah$$

$$g = bf - ec$$

$$h = cd - af$$

$$i = ae - bd$$

$$ad + be + cf = 0$$

$$dg + eh + fi = 0$$

$$ag + bh + ci = 0$$

Once the elements of $[T]$ are known, the unknowns are found as follows:

First, let

$$j_1 = t_{11}t_{32} - t_{12}t_{31}$$

$$j_1 = ((k_1a + u_0g)/r) (h/r) - ((k_1b + u_0h)/r) (g/r)$$

$$j_1 = (ah - bf) (k_1/r^2)$$

Using the constraint that $f = bg - ah$,

$$j_1 = -k_1f/r^2$$

Likewise, let

$$j_2 = t_{21}t_{32} - t_{22}t_{31}$$

$$j_2 = ((k_2d + v_0g)/r) (h/r) - ((k_2e + v_0h)/r) (g/r)$$

$$j_2 = (hd - eg) (k_2/r^2)$$

Using the constraint that $c = hd - eg$,

$$j_2 = k_2 c / r^2$$

Remember that since t_{ij} are known, j_1 and j_2 are known.

Also,

$$(t_{31})^2 + (t_{32})^2 = (g/r)^2 + (h/r)^2 = (g^2 + h^2)/r^2$$

Expressing the constraints as,

$$g^2 + h^2 + i^2 = c^2 + f^2 + i^2 = 1$$

gives,

$$g^2 + h^2 = c^2 + f^2$$

Substituting this into the expression for

$$(t_{31})^2 + (t_{32})^2$$

gives:

$$(t_{31})^2 + (t_{32})^2 = (f^2 + c^2)/r^2 = r^2(f^2 + c^2)/r^4$$

Now substitute for f and c values using j_1 and j_2 :

$$(t_{31})^2 + (t_{32})^2 = ((j_1/k_1)^2 + (j_2/k_2)^2) r^2$$

Because t_{31} , t_{32} , j_1 , j_2 , k_1 , and k_2 are known, this equation can be solved for r^2 and r (assume that r is positive if the image is in front of the camera).

Then, since $f^2 + c^2 = 1 - i^2$, substitute and solve $(t_{31})^2 + (t_{32})^2 = (1 - i^2)/r^2$ for i^2 .

i will be positive if the cross product of the U and V axes is in the same direction as the Z axis and negative if in

the opposite direction. This will depend on the way that the system coordinate system is initially established.

Now, knowing r and i , other values follow quickly:

$$t_{31} = g/r, \quad g = t_{31}r$$

$$t_{32} = h/r, \quad h = t_{32}r$$

$$j_1 = k_1 f/r^2, \quad f = j_1 r^2/k_1$$

$$j_2 = k_2 c/r^2, \quad c = j_2 r^2/k_2$$

Next, look at

$$\begin{aligned} t_{11}t_{31} + t_{12}t_{32} &= ((k_1 a + u_0 g)/r) (g/r) + \\ &\quad ((k_1 b + u_0 h)/r) (h/r) \\ &= (k_1/r^2) (ag + bh) + \\ &\quad (u_0/r^2) (g^2 + h^2) \end{aligned}$$

From the constraint, $ag + bh + ci = 0$, make the substitution:

$$ag + bh = -ci$$

$$\text{Then, } t_{11}t_{31} + t_{12}t_{32} = -(k_1 ci/r^2) + u_0(g^2 + h^2)/r^2$$

$$t_{11}t_{31} + t_{12}t_{32} + k_1 ci/r^2 = u_0(g^2 + h^2)/r^2$$

and u_0 can be determined:

$$u_0 = (r^2/(g^2 + h^2)) (t_{11}t_{31} + t_{12}t_{32} + k_1 ci/r^2)$$

Now, look at

$$\begin{aligned} t_{21}t_{31} + t_{22}t_{32} &= ((k_2 d + v_0 g)/r) (g/r) + \\ &\quad ((k_2 e + v_0 h)/r) (h/r) \end{aligned}$$

$$= (k_2/r^2) (dg + eh) + \\ (v_o/r^2) (g^2 + h^2)$$

From the constraint, $dg + eh + fi = 0$, make the substitution:

$$dg + eh = -fi$$

$$\text{Then, } t_{21}t_{31} + t_{22}t_{32} = -(k_2fi/r^2) + \\ (v_o/r^2) (g^2 + h^2)$$

and v_o can be determined:

$$v_o = (r^2/(g^2 + h^2)) (t_{21}t_{31} + t_{22}t_{32} + k_2fi/r^2)$$

Now, solve for a, b, p, d, e , and q as follows:

$$\text{since } t_{11} = (k_1a + u_og)/r,$$

$$a = (rt_{11} - u_og)/k_1$$

$$\text{since } t_{12} = (k_1b + u_oh)/r,$$

$$b = (rt_{12} - u_oh)/k_1$$

$$\text{since } t_{14} = (k_1p + u_or)/r,$$

$$p = (r/k_1) (t_{14} - u_o)$$

$$\text{since } t_{21} = (k_2d + v_og)/r,$$

$$d = (rt_{21} - v_og)/k_2$$

$$\text{since } t_{22} = (k_2e + v_oh)/r,$$

$$e = (rt_{22} - v_oh)/k_2$$

$$\text{since } t_{24} = (k_2q + v_or)/r,$$

$$q = (r/k_2) (t_{24} - v_o)$$

Because $a, b, c, d, e, f, g, h, i, p, q$, and r are now

known, it is possible to use them to determine X_s , Y_s , Z_s , l , m , and n .

To solve for X_s , Y_s , and Z_s solve equations:

$$X_s = -ap - dq - gr$$

$$Y_s = -bp - eq - hr$$

$$Z_s = -cp - fq - ir$$

To find l , m , and n first use $f = \sin m$ and solve for m then use this value in the equations for e and i to obtain l and n :

$$m = \sin^{-1}(f)$$

$$l = \cos^{-1}(e/\cos(m))$$

$$n = \cos^{-1}(i/\cos(m))$$

The above presentation shows that by knowing the focal length and scaling of the camera, one can find the position and orientation of the camera if the target has four features that can be identified and located. Other methods of determining camera location solve the general approach by identifying more target spots and using least squares methods of getting a best fit that satisfies the transformation equations and the constraint equations [Fischler, 1980], [Huang, 1981], [Lowe, 1980], [Sutherland, 1974].

The method described in this paper relies heavily on accurate measurements of u_i and v_i . It is generally more

accurate to use k_1 and k_2 as known values (vice u_0 and v_0) in solving the problem because k_1 and k_2 are typically more accurately found when calibrating the camera. Because this solution method is not exact, the results may need some conditioning if the measurements of u_1 and v_1 are not accurate enough for the desired application. However, this closed form solution to the problem has the advantage of using a simple target design and requiring little computing power to obtain the solution.

Of course, using this method depends on the ability to find the elements of $[T]$. A description of how these elements can be found follows.

2.2 Determining the Transformation Matrix

The coordinates in the above matrices have been expressed using homogeneous coordinate systems. In a homogeneous coordinate system, the three dimensional point x,y,z is represented by the four-tuple w_x, w_y, w_z, w and the two dimensional point u,v is represented by the three-tuple w_u, w_v, w . After using the homogeneous coordinates, the three and two dimensional coordinates are obtained by dividing all components by the last component, w . So, although the perspective transformation has been written as:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{14} \\ t_{21} & t_{22} & t_{24} \\ t_{31} & t_{32} & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

in practice, after performing the matrix multiplication, the first two rows are divided by the third. This gives the following expressions for u_i and v_i :

$$u_i = (t_{11}x_i + t_{12}y_i + t_{14}) / (t_{31}x_i + t_{32}y_i + 1)$$

$$v_i = (t_{21}x_i + t_{22}y_i + t_{24}) / (t_{31}x_i + t_{32}y_i + 1)$$

where $i = 1, 2, 3, 4$

This yields eight equations to find the eight unknowns $t_{11}, t_{12}, t_{14}, t_{21}, t_{22}, t_{24}, t_{31}, t_{32}$:

$$u_1 + u_1x_1t_{31} + u_1y_1t_{32} - x_1t_{11} - y_1t_{12} - t_{14} = 0$$

$$u_2 + u_2x_2t_{31} + u_2y_2t_{32} - x_2t_{11} - y_2t_{12} - t_{14} = 0$$

$$u_3 + u_3x_3t_{31} + u_3y_3t_{32} - x_3t_{11} - y_3t_{12} - t_{14} = 0$$

$$u_4 + u_4x_4t_{31} + u_4y_4t_{32} - x_4t_{11} - y_4t_{12} - t_{14} = 0$$

$$v_1 + v_1x_1t_{31} + v_1y_1t_{32} - x_1t_{21} - y_1t_{22} - t_{24} = 0$$

$$v_2 + v_2x_2t_{31} + v_2y_2t_{32} - x_2t_{21} - y_2t_{22} - t_{24} = 0$$

$$v_3 + v_3x_3t_{31} + v_3y_3t_{32} - x_3t_{21} - y_3t_{22} - t_{24} = 0$$

$$v_4 + v_4x_4t_{31} + v_4y_4t_{32} - x_4t_{21} - y_4t_{22} - t_{24} = 0$$

The above equations can be written in matrix form as:

$$\begin{array}{c}
 - \\
 \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}
 \end{array}
 =
 \begin{bmatrix}
 -x_1 & -y_1 & -1 & 0 & 0 & 0 & x_1 u_1 & y_1 u_1 \\
 -x_2 & -y_2 & -1 & 0 & 0 & 0 & x_2 u_2 & y_2 u_2 \\
 -x_3 & -y_3 & -1 & 0 & 0 & 0 & x_3 u_3 & y_3 u_3 \\
 -x_4 & -y_4 & -1 & 0 & 0 & 0 & x_4 u_4 & y_4 u_4 \\
 0 & 0 & 0 & -x_1 & -y_1 & -1 & x_1 v_1 & y_1 v_1 \\
 0 & 0 & 0 & -x_2 & -y_2 & -1 & x_2 v_2 & y_2 v_2 \\
 0 & 0 & 0 & -x_3 & -y_3 & -1 & x_3 v_3 & y_3 v_3 \\
 0 & 0 & 0 & -x_4 & -y_4 & -1 & x_4 v_4 & y_4 v_4
 \end{bmatrix}
 \begin{bmatrix} t_{11} \\ t_{12} \\ t_{14} \\ t_{21} \\ t_{22} \\ t_{24} \\ t_{31} \\ t_{32} \end{bmatrix}$$

Chapter 3 Use of Prepared, Passive Targets

3.1 Target Design Criteria

Several methods have been developed to determine the position of a camera and track the motion of a camera relative to a landmark or target. Some methods involve "teaching" the tracking system some general shapes it can identify and then determining its position relative to the shapes by use of stereo vision or some other method [Nasr, 1988], [Sugihara, 1988]. A second approach is to construct a target of spots and, knowing the location of the spots on the target, use perspective transformation analysis to determine camera position.

All of the applications discussed in this paper involve operating and controlling an ROV in an underwater environment. The terrain and possible landmarks are varied depending on the location and depth. In fact, in many cases, the shapes present may not be known (if they exist at all) until the ROV is at the site. For this reason, "teaching" the tracking system some general shapes or landmarks may be impractical (except in a limited number of applications). Also, the algorithms used for determining camera position relative to landmarks are typically numerically based and

involve either long computing times or large amounts of computer memory space.

A simpler, more direct method uses perspective transformations to orient and position a target consisting of a known pattern of points. As explained in Chapter 2, a closed form solution of the problem is obtained by using four target spots and information about the camera. Once the actual spot coordinates and the image plane spot coordinates are known, the transformation matrix can be found and its elements used to solve for the camera position and orientation.

Given that a prepared target will be used, it must now be decided whether the target will be "active" or "passive". An active target uses a power source to illuminate the target spots so that the camera can detect them. For example, some tracking systems use infrared light emitting diodes as the target spots. These IRED spots act as beacons to the camera; the camera detects the location of the infrared light and the transformation matrix can be determined. Passive targets use ambient or reflected light to illuminate the spots. For example, the target spots can be made of a reflective material and light from lights or a strobe will be reflected from the spots into the camera lens where the image plane spot

coordinates can be determined.

For underwater applications, where light is absorbed over short distances and where it is difficult to keep a target powered for long periods of time, an active target may be impractical. Whereas the JASON ROV is powered from shipboard generators via the ARGO, an active target would require its own power supply or a link to the ship. Any battery powered active target would need occasional battery replacement and that replacement might be needed at very inconvenient times. An active target, once activated, would emit its signal even if the camera were too far away to detect the signal and so the signal (and its power) would be wasted. A passive target would not require any type of power supply and would send out a signal only if the camera were close enough to illuminate the target. For these reasons, a passive target was designed and used in the measurements presented in this paper.

The target used in the optical tracking system for JASON was composed of four reflective spots. The spots were placed in general position on a flat ($z = 0$), non-reflective surface. The x, y coordinates of these spots make up the x_1, y_1 matrix used in Chapter 2 to determine the transformation matrix. In the solution method, these values will not change.

As one of JASON's video cameras is pointed toward the

target, the light from the camera light reflects off the spots and into the camera lens where the image of the target is projected onto the image plane. The ISCAN video tracking system detects the location of the target spots as they appear on the image plane. These values give the u_i, v_i matrix of the solution method. As JASON moves toward and away from the target and as JASON changes its pitch, roll, and yaw, the elements of the u_i, v_i matrix change. Solving for the new transformation matrix that is produced each time JASON changes its position and orientation gives the current values of X_s, Y_s, Z_s and l, m, n . These position and orientation values can then be used in the JASON control functions to navigate, keep station, or dock.

3.2 Target Spot Placement

The target spots need to be placed on the target so that they will be inside the camera field of view at the point of closest approach. The spots need also to be placed far enough apart so that they are distinguishable from each other at long ranges. They must be far enough away from each other that the spot tracking system can detect each spot individually.

If we assume that the ROV will be .25 meters away from the target at the point of closest approach, the target spots

should fill the image plane (this will give the maximum spot spread at long range). Looking straight at the target at this point of closest approach, the maximum spread of the spots can be determined (this will also depend on the camera focal length).

Because the spots are positioned generally, (non-symmetric) this allows the spots to be placed near the edges of this window. Figure 8 gives a possible spot positioning.

Chapter 4 Description of the ISCAN Video Tracking System

The ISCAN RK-446 Video Tracking System is a real time digital image processor that tracks and measures contrasting targets in the field of view of a video camera. The system determines the center of a target spot to within one-half pixel giving the horizontal and vertical coordinates of the target spot as they appear on the image plane of the camera. The system updates the horizontal and vertical position measurements every 16 msec (at a frequency of 60 Hz) and gives this information in a computer compatible, parallel format. The image processor uses an artifact elimination algorithm that rejects noise or clutter. The RK-446 divides the image plane into a grid of 511 pixels horizontally and 255 pixels vertically.

Chapter 5 Testing

5.1 Test Design

To test the above described analytical method, a test target of four spots in general position was designed. The target spots were precision placed and drilled to be within one mil (within one thousandth of an inch). The target spots were placed so that the target Y coordinate axis was in the middle of the target plate and the X coordinate axis was 5 cm above the base plate. This positioning gives the x_i and y_i values given in appendix A.

The target plate was placed perpendicular to the base plate. The base plate was made of plexiglass and had five holes drilled at the locations shown in figure 9. Pegs of various heights were placed in the holes and the camera was then sighted along these so that a known camera orientation could be determined. For example, by placing pegs 5 cm high into holes A and B and sighting along these pegs, the camera was at a height of $Y=0$ and at a pan angle of 0° (relative to the target coordinate system). Pegs in holes A and C gave a pan angle of 30° . Pegs in holes A and D gave a pan angle of 45° . Pegs in holes A and E gave a pan angle of 60° . By placing a 10 cm peg in hole B, a tilt angle of 25.7° was

obtained. A 14 cm peg in hole B gave a tilt angle of 40.87° .

The target was placed on a table over which calibrated paper was placed. Using the RK-446 Video Tracking System, the values of u_i and v_i were recorded as the camera was placed at various positions and orientations. The coordinates and results of these measurements are listed in the appendices.

5.2 Camera Calibration

Prior to complete testing, it is necessary to calibrate the camera so that the horizontal and vertical scaling factors, k_1 and k_2 can be determined.

The camera was calibrated by placing the four target spots in two parallel planes. This gave eight image plane spot measurements. The target was placed in front of the camera and the image plane coordinates of the target spots were recorded. The target was then moved a known distance along the z axis. At the next target location, the image plane spot coordinates were again recorded. Knowing the x_i , y_i , z_i coordinates of the target spots and the u_i , v_i measurements of the eight spots gave a total of sixteen equations to find the eleven unknowns of the transformation matrix (elements t_{11} through t_{34}). A least-squares approach was used to find the transformation matrix elements. Once these elements were

known, the system constraint equations were used to solve for k_1 and k_2 . (In much the same way as described in Chapter 2 but by using the full transformation matrix since no camera parameters were yet known.)

This calibration was done for the three lenses to be tested: a 4.8mm focal length lens, an 8mm focal length lens, and a 12mm focal length lens. The calibration results for k_1 and k_2 for each lens are given in appendices A, B, and C respectively.

5.3 Test Results

The first system tests were performed using the test target and a Sony DXC-101 ccd video camera in air. The target spot coordinates, known camera location and orientation, image plane measurements, and algorithm results are listed in appendices A, B, and C for the 4.8mm, 8mm, and 12mm lenses respectively.

There are several sources of possible error. First, the solution method assumes that the camera behaves as a pin-hole camera. The 4.8mm lens is probably too short for this to be true. For the longer lenses, the error due to non-pin-hole behavior is inconsequential. Errors due to imperfect calibration are also possible (this is probably why the 8mm

lens results are not as good as those of the 12mm lens). Third, the ISCAN Video Tracking System is accurate to within one-half pixel when finding the target spot centers. Any error in finding the spot centers would result in incorrect image plane coordinate measurements. The method seems to be largely unaffected by such errors when the longer focal length lenses are used. Inputting half pixel changes in the measurements did slightly change the X_s and/or Y_s results but had a smaller affect on the Z_s result. However, doing the same for the 4.8mm lens had dramatic effects. For example, a non-symmetric half-pixel change in the u_1 and v_1 measurements changed the X_s and/or Y_s results by as much as 50% (but again had little affect on the Z_s result). This is probably due to the less pin-hole behavior of the 4.8mm lens. The consistently largest and most important source of error for the system, independent of the lens used, results when the system is used underwater and is subjected to the effects of refraction.

Chapter 6 Underwater Effects

To this point, all analysis and experimentation has involved the camera, target, and tracking system in an air environment. When the system is placed in the water, optical effects can dramatically affect the system's performance.

For the target, the underwater environment creates no difficulties. The reflectance of the spots will be the same as in air and will not change with depth.

For the camera, there are the problems of turbulence, suspended particles, and refraction.

Turbulence in the water gives rise to inhomogeneities along the optical paths of the camera. These patches of water have different optical properties than surrounding water and cause optical distortion. However, these effects are small and are generally a factor only in long range, high resolution applications.

Suspended particles scatter the light and create a type of veil between the target and the camera. This veil of scattered light reduces the contrast between the light and dark of the target. However, by designing target spots that are initially in sharp contrast with the target plate, this effect can be ignored.

Refraction is the bending of a ray of light as it passes through materials of different optical properties. The cameras on JASON are housed inside cases that have a flat window through which the cameras look. This gives three interfaces: the water/window interface, the window/air interface, and the air/lens interface. This system can be modelled as consisting of two components, the window producing an image and the air/lens focusing that image onto the image plane. Because the index of refraction in water is approximately 1.33 times greater than the index of refraction in air, the distance from the lens to the target along the line of sight will appear to be 1.33 times shorter than it really is. Also, because of the bending of the light rays at the air/lens interface, lateral distances will seem different than they actually are. For example, a horizontal field of view in air of 77.3° will appear as a 56° field of view in water. When the camera looks at the target and the image plane coordinates of the target spots are measured, they will appear farther apart than they should be.

Because the decreased field of view and the shortening of length along the line of sight are independent of distance, it should be possible to implement corrections in the algorithm.

It is worth mentioning that the water index of refraction

is a function of depth, salinity, and turbidity and can change by as much as two percent. Because the cameras used in determining position and orientation have fixed focal lengths, this variability in the index of refraction may affect system accuracy (by slightly shifting the focal length). However, since the index of refraction error is less than the system error, it is probably not necessary to account for these changes in the algorithm.

In using perspective transformations, there are two options to pursue.

First, because refraction moves the image plane spots farther apart than they would appear in air, the spots can be scaled back to where they would appear without refraction. However, the distance the spots should be scaled depends on the incident angle of the light which, in turn, depends on the distance to the target. To be completely accurate, the u_i and v_i measurements should be scaled by 1.33 (the ratio of the index of refraction in water and the index of refraction in air) and the perspective transformation performed. The factor of 1.33 should then be re-evaluated by using the computed distance to the target. This new factor (slightly larger than 1.33) should then be used to re-scale the u_i and v_i measurements and a new perspective transformation should

be performed. Using the newly computed distance to the target, the factor of 1.33 should again be revised and the process repeated until the result settles on the actual values.

However, this process takes us away from our goal of avoiding any time consuming numerical methods (like iteration). By analyzing the refraction geometry, one finds that if the u_i and v_i measurements are scaled by 1.33 and the perspective transformation performed, the result will be within allowable error domains. For the test target, doing a one-time scaling of 1.33 introduces an error of less than three percent for an 8mm lens and an error of less than one percent for a 12mm lens. For a 4.8mm lens, the error is larger and the scaling introduces the instability previously described.

A second method of correcting for water effects is to use a corrective dome as the window on the camera housing. A corrective dome is designed for a specific focal length lens and a specific camera housing. The dome corrects the refraction so that the image appears as it would in air. However, corrective domes tend to shrink and slide in high pressure environments (which causes some image distortion) and can be quite expensive to produce. In general, corrective

domes work well for wide angle lenses (air fields of view greater than eighty degrees) and would be the best approach for the 4.8mm lens.

Chapter 7 Conclusions

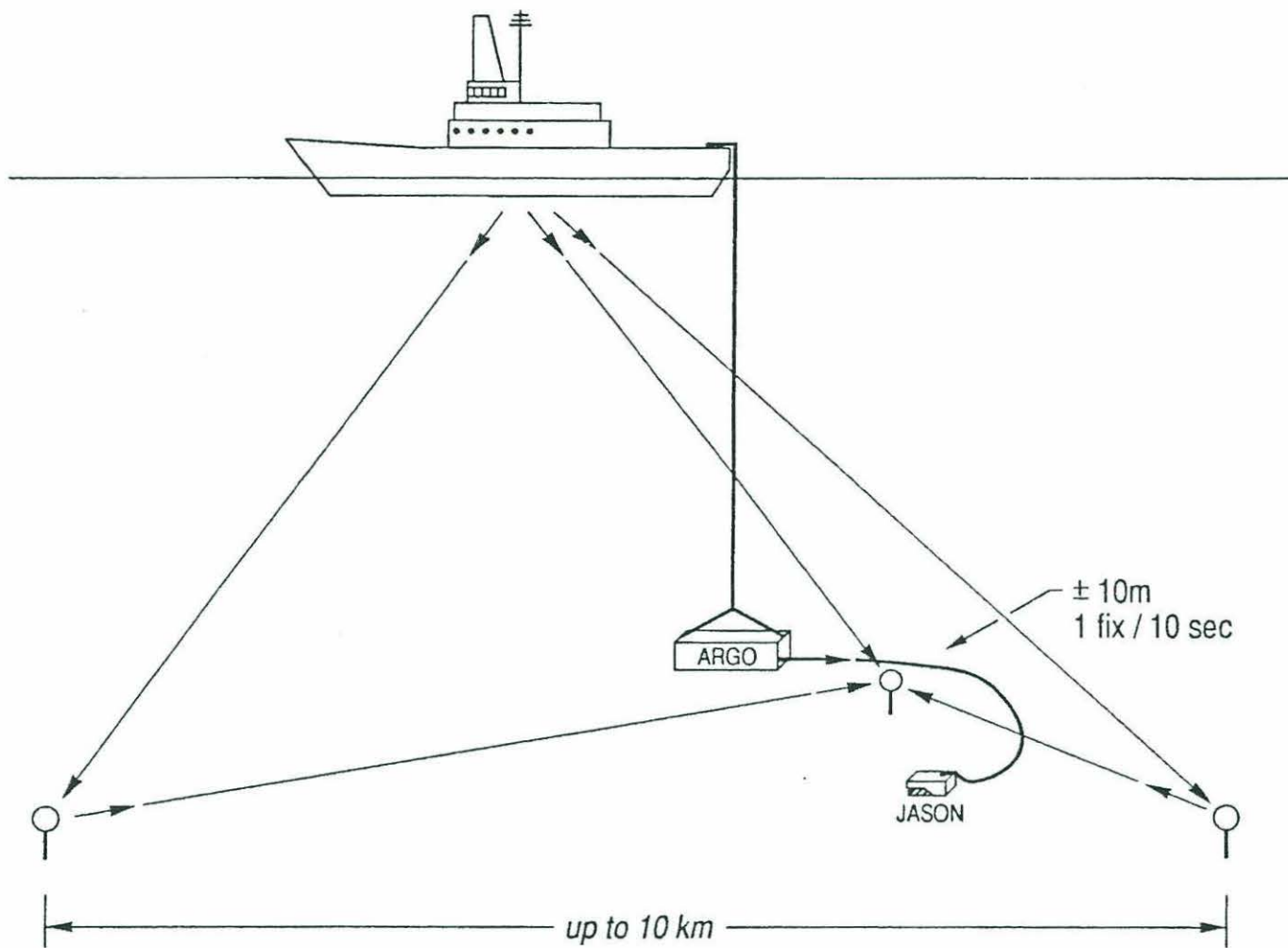
The method described in this paper seems to work well for camera lenses with focal lengths of at least 8mm. Underwater effects can, to within an acceptable degree of error, be compensated for by scaling the u_i and v_i measurements toward the image plane center by a factor of 1.33.

The method does not work as well for shorter focal length lenses (4.8mm in particular) although the computed range values are reasonable and usable. Errors introduced by using a 4.8mm lens will be increased when underwater but can be improved to in-air values by using a compensating dome as the window on the camera housing.

Errors in the u_i and v_i measurements affect the X_s , Y_s , and orientation calculations but have a much smaller effect on the Z_s calculation.

The method is non-numerical and should be easy to implement. The U,V measurements from the ISCAN Video Tracking System can be output to a JASON computer (via a parallel I/O interface) where real-time ROV position can be calculated. This position can then be used in JASON control functions.

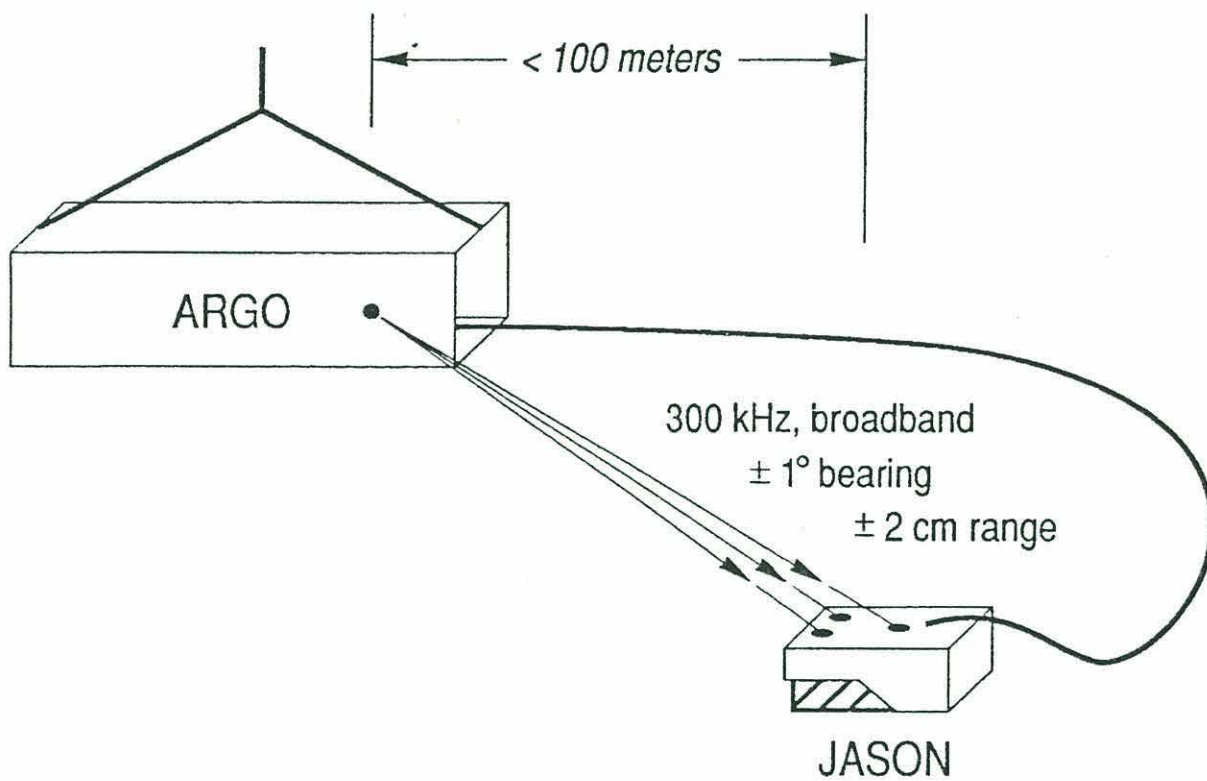
Low Frequency Long Baseline Narrow Band Pulses ~10 kHz



ARGONAUT/JASON Navigation Referenced to Ship

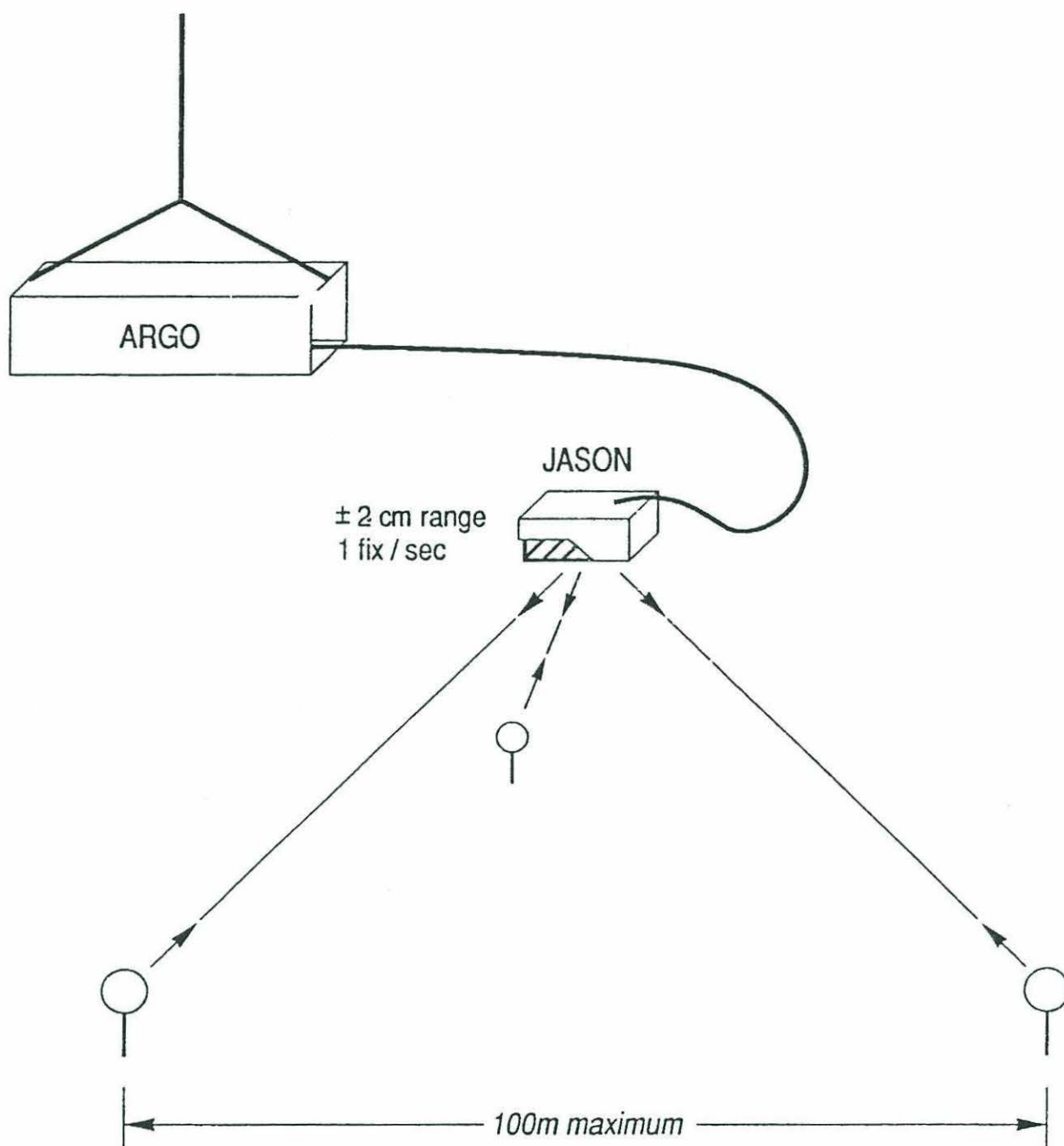
Figure 1

High Frequency, Relative to ARGO



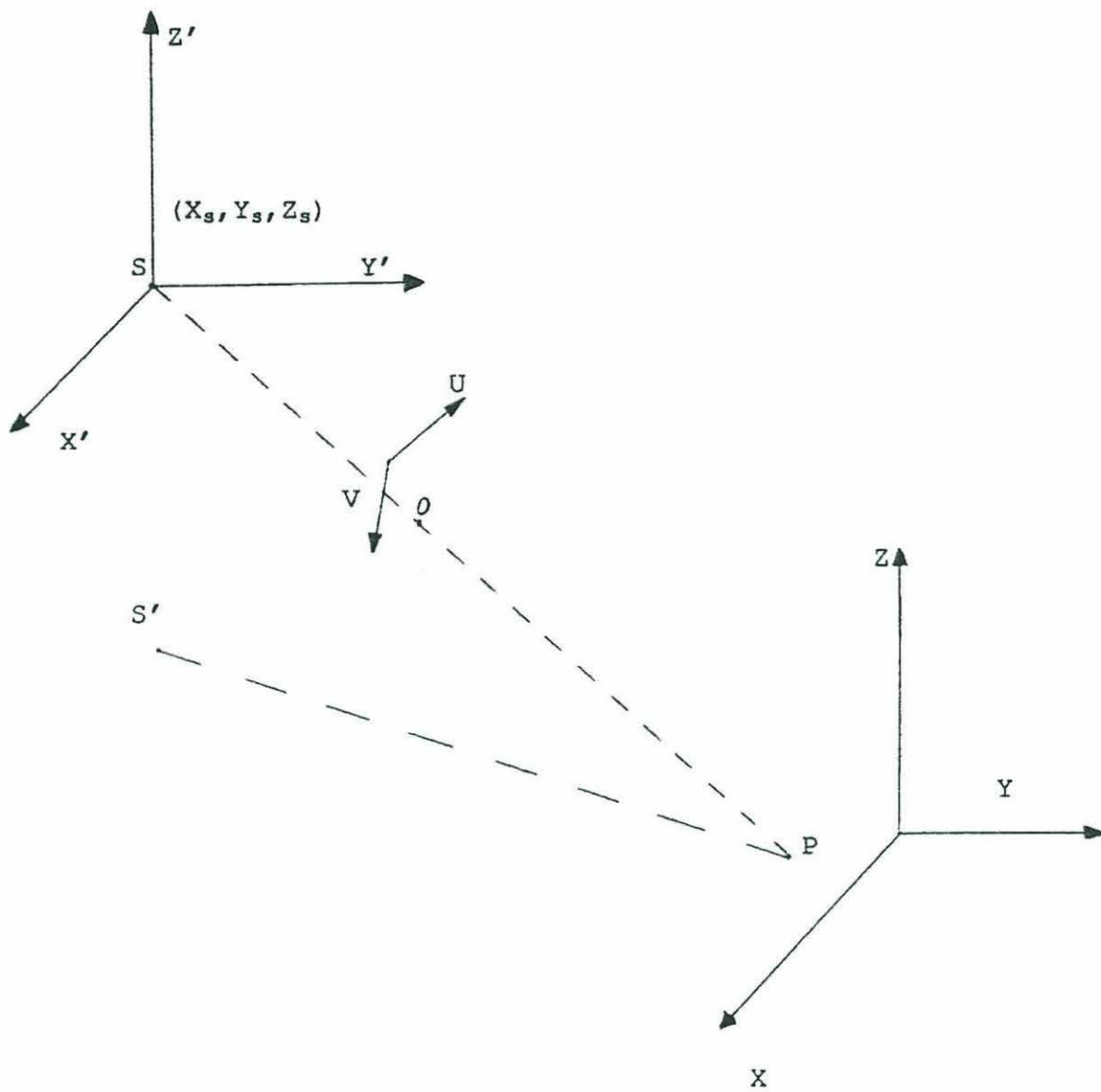
JASON Navigation Referenced to ARGO

Figure 2



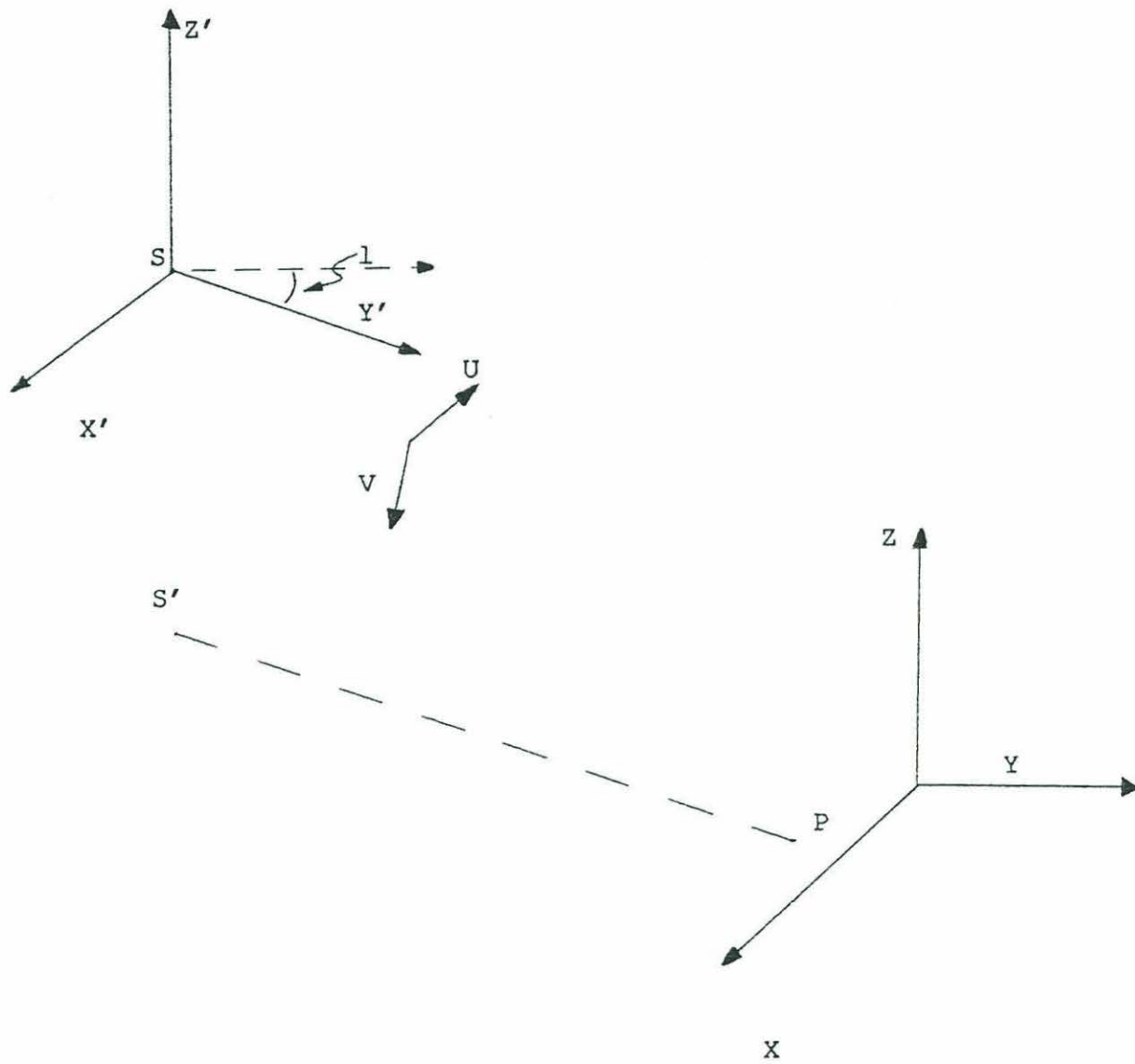
JASON Navigation Referenced to Seafloor

Figure 3



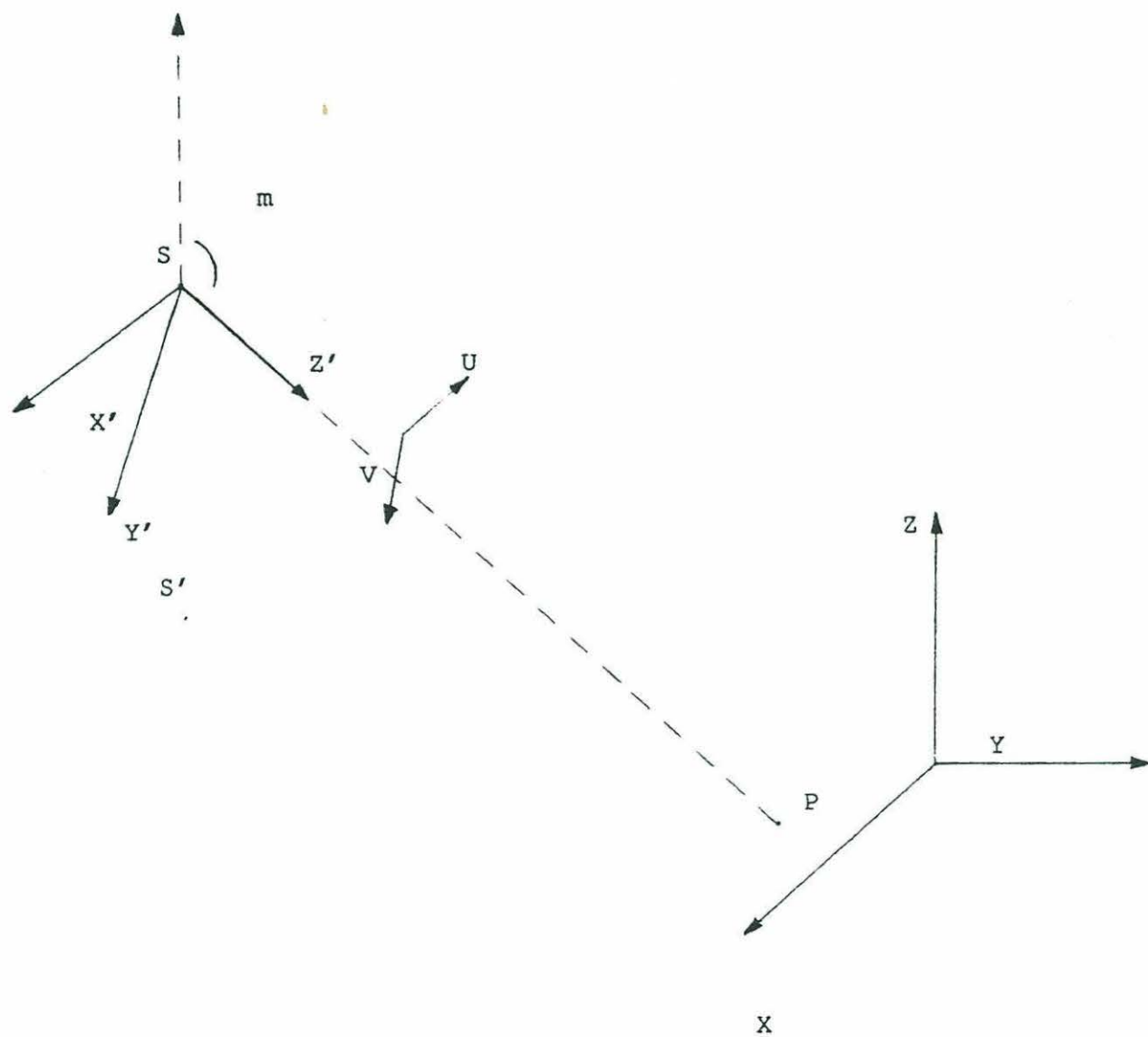
Displacement of Camera Coordinate System

Figure 4



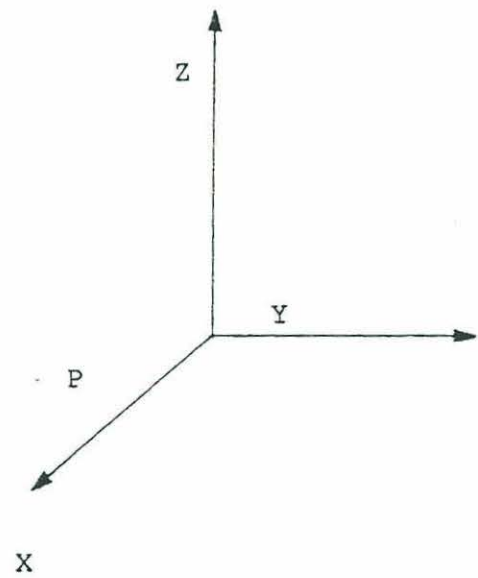
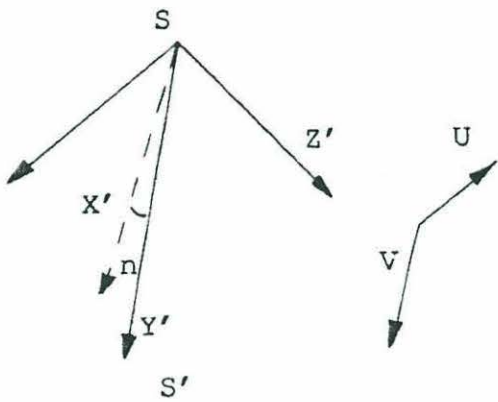
Pan Angle Rotation

Figure 5



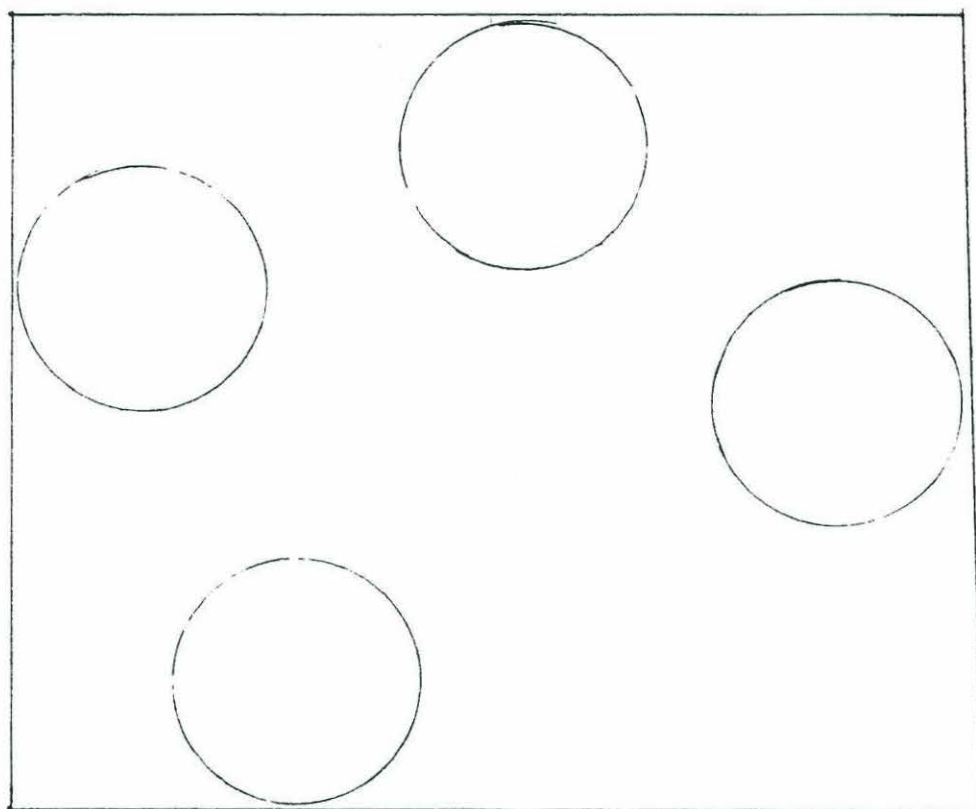
Tilt Angle Rotation

Figure 6



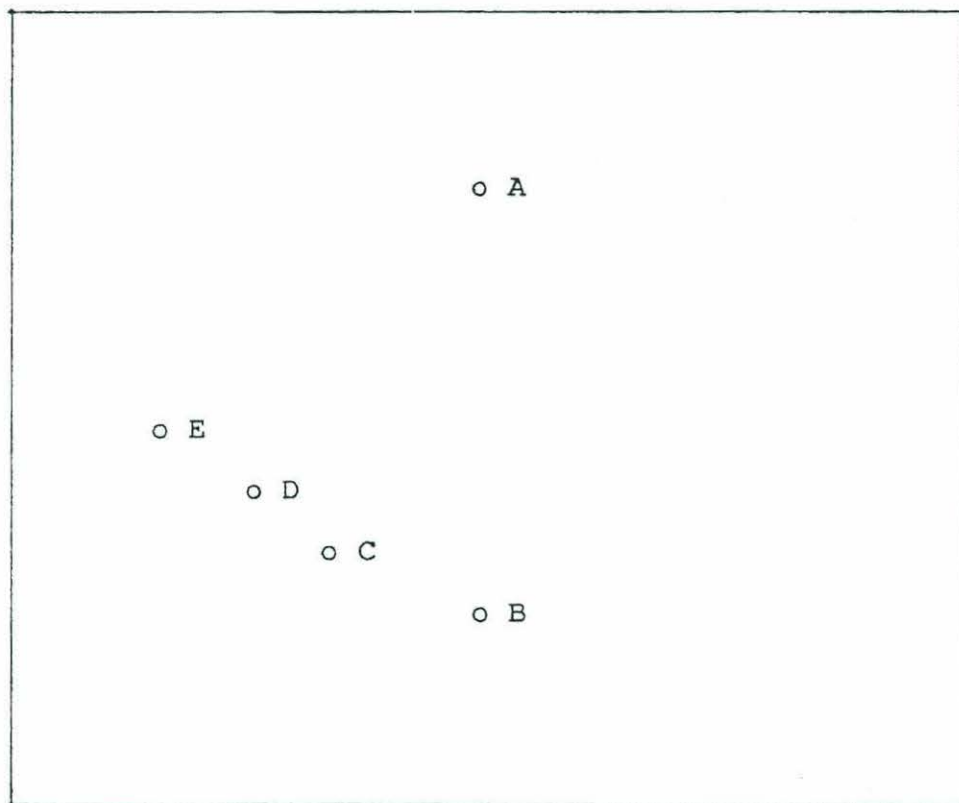
Swing Angle Rotation

Figure 7



Possible Target Spot Location

Figure 8



Test Target Base Plate

Figure 9

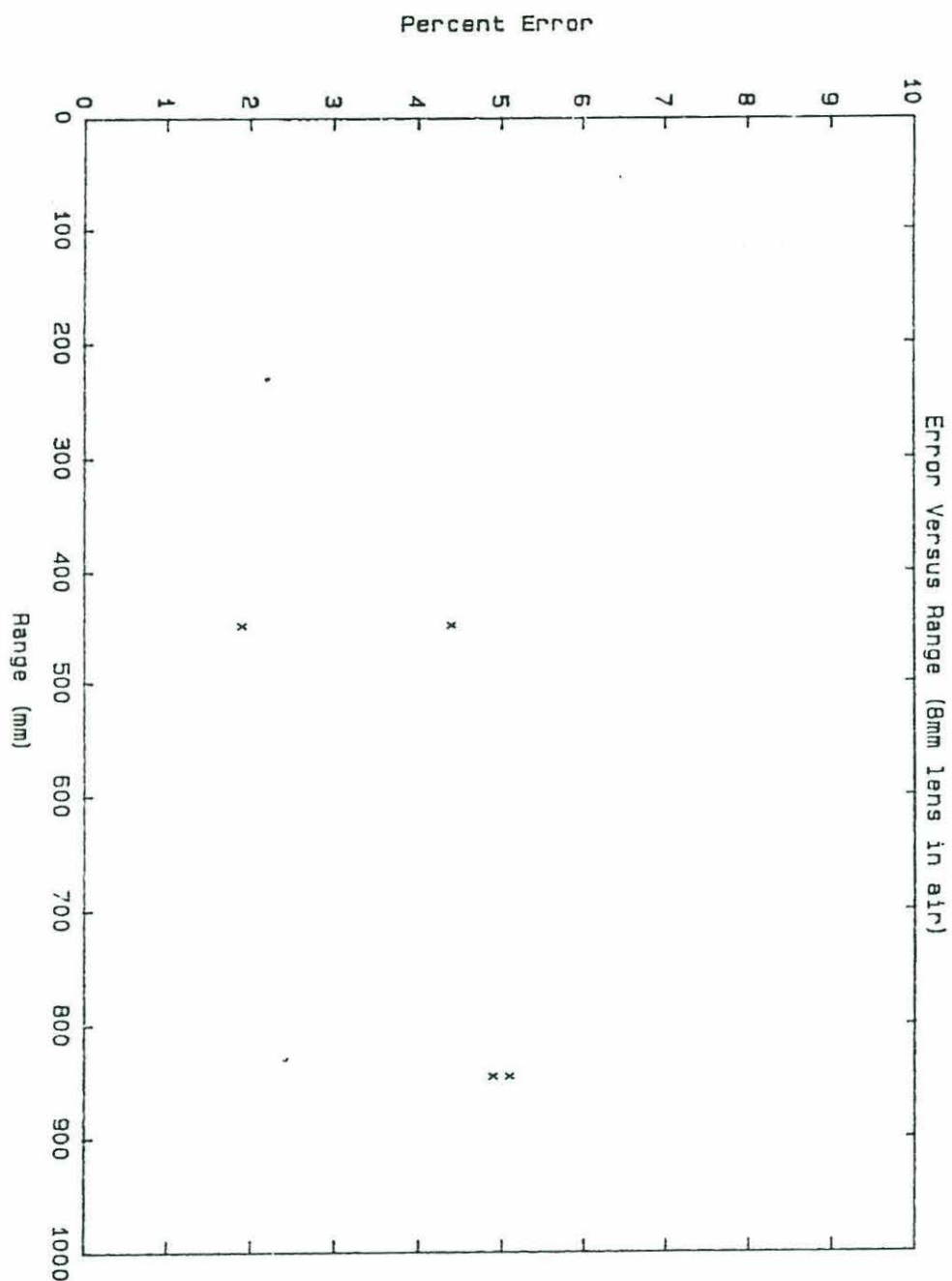


Figure 10

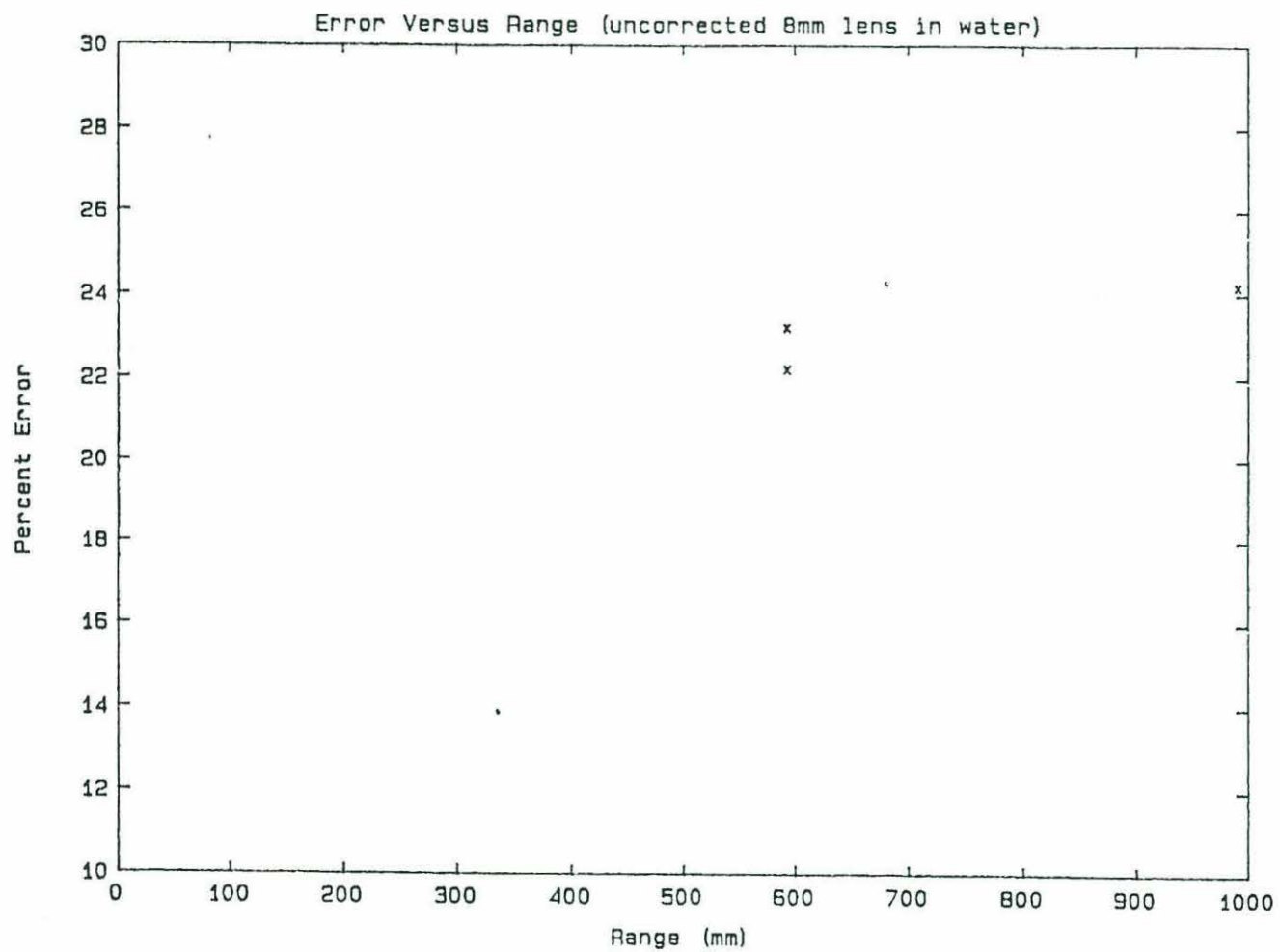


Figure 11

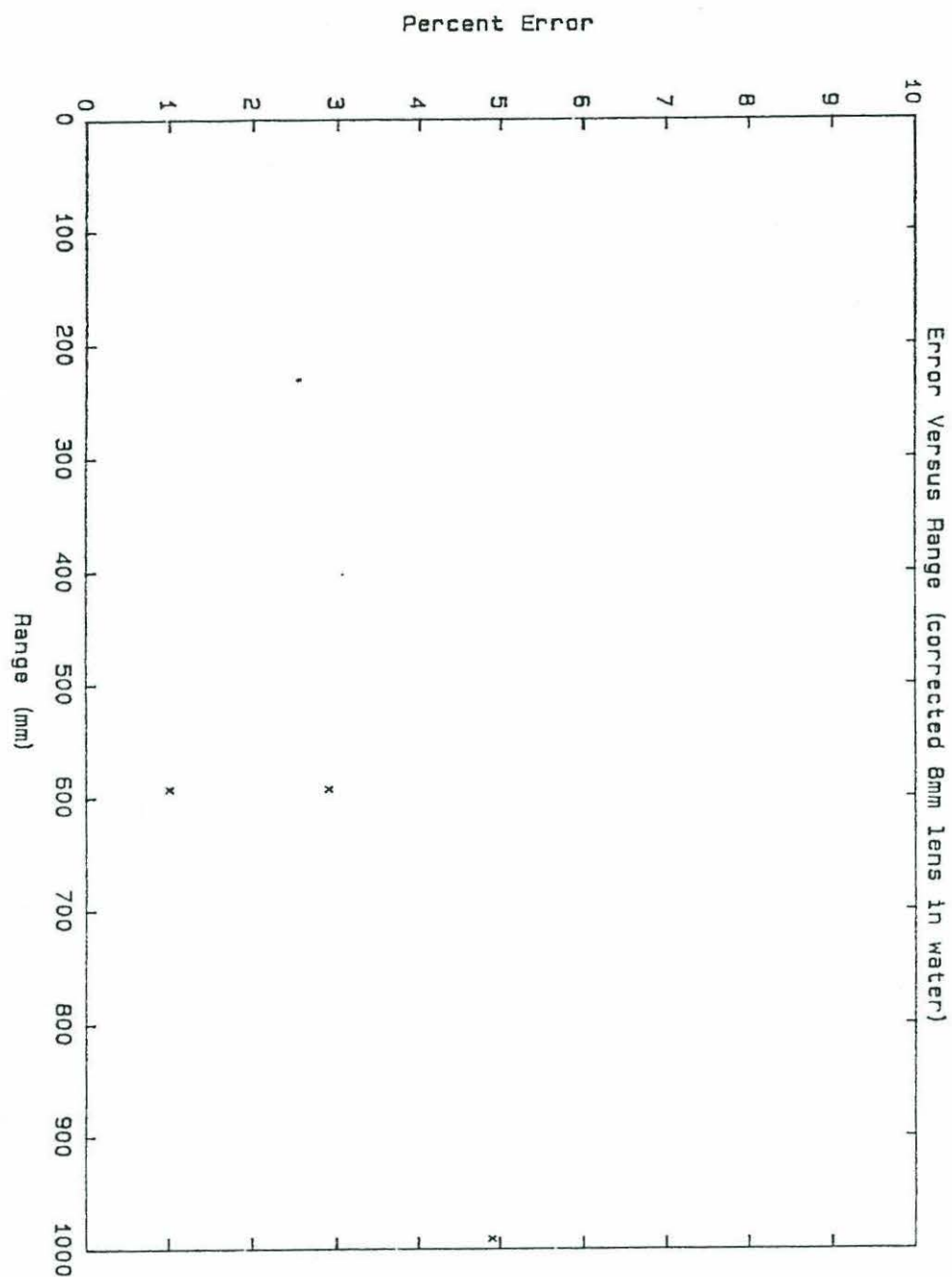


Figure 12

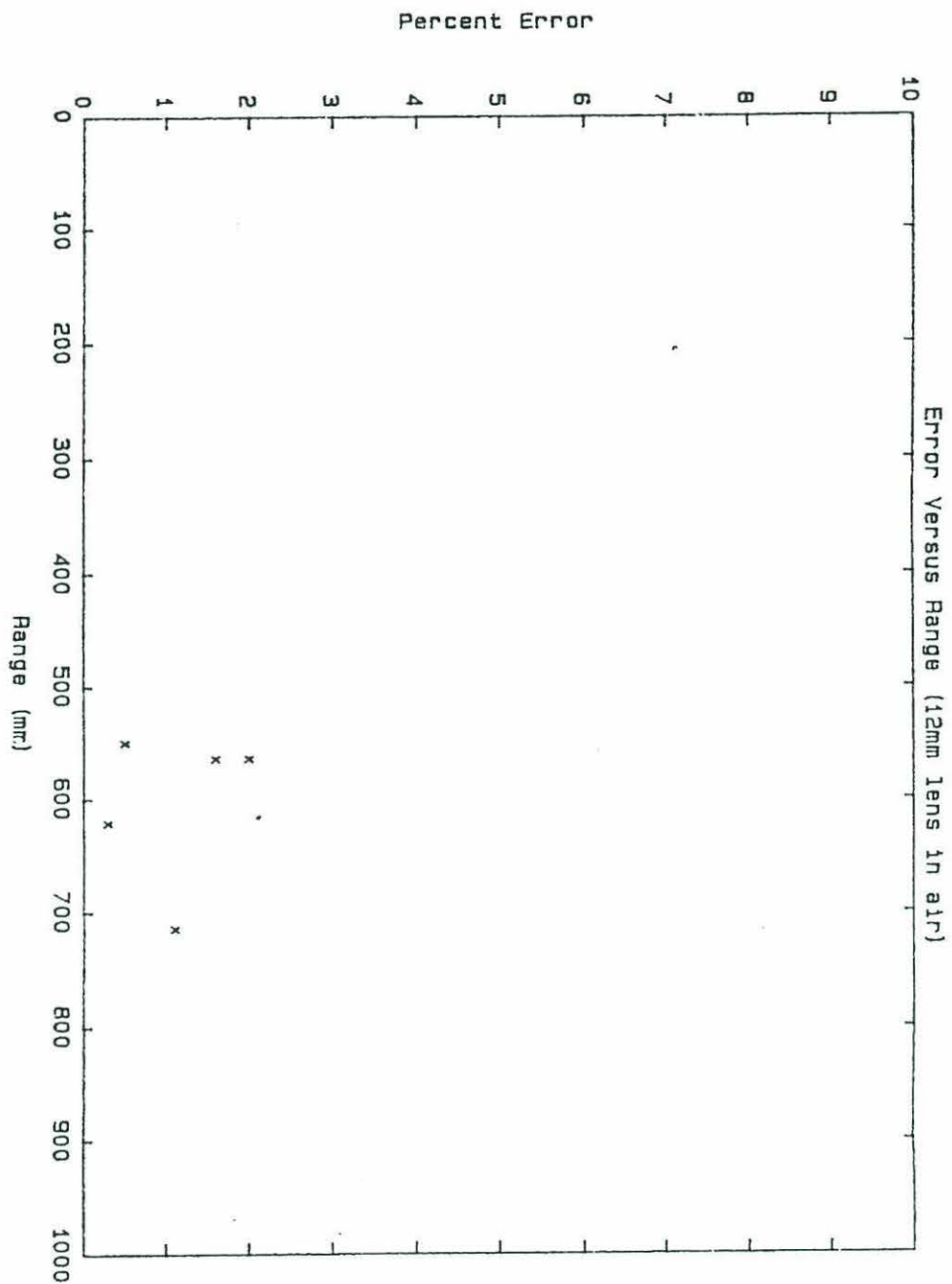


Figure 13

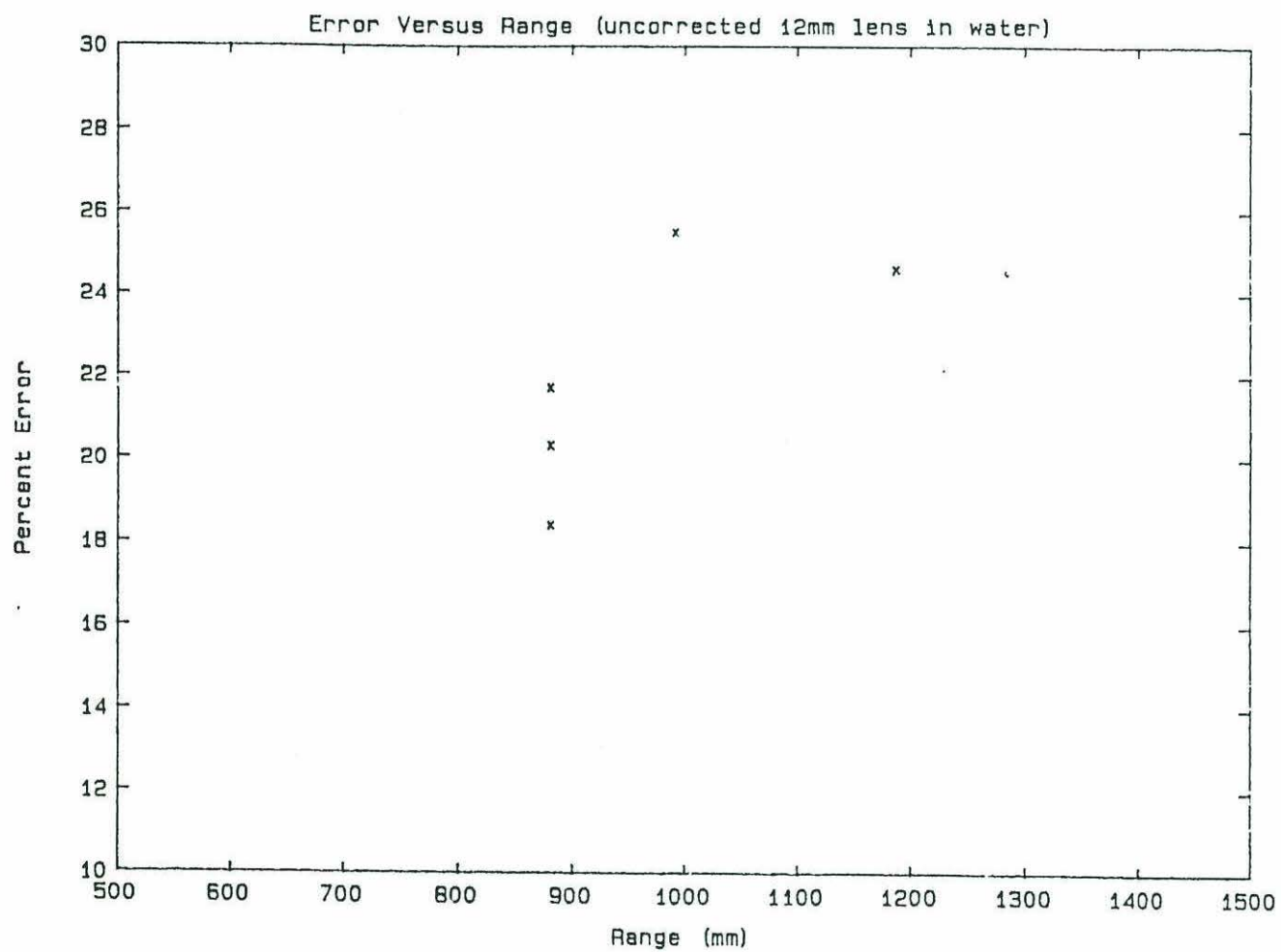


Figure 14

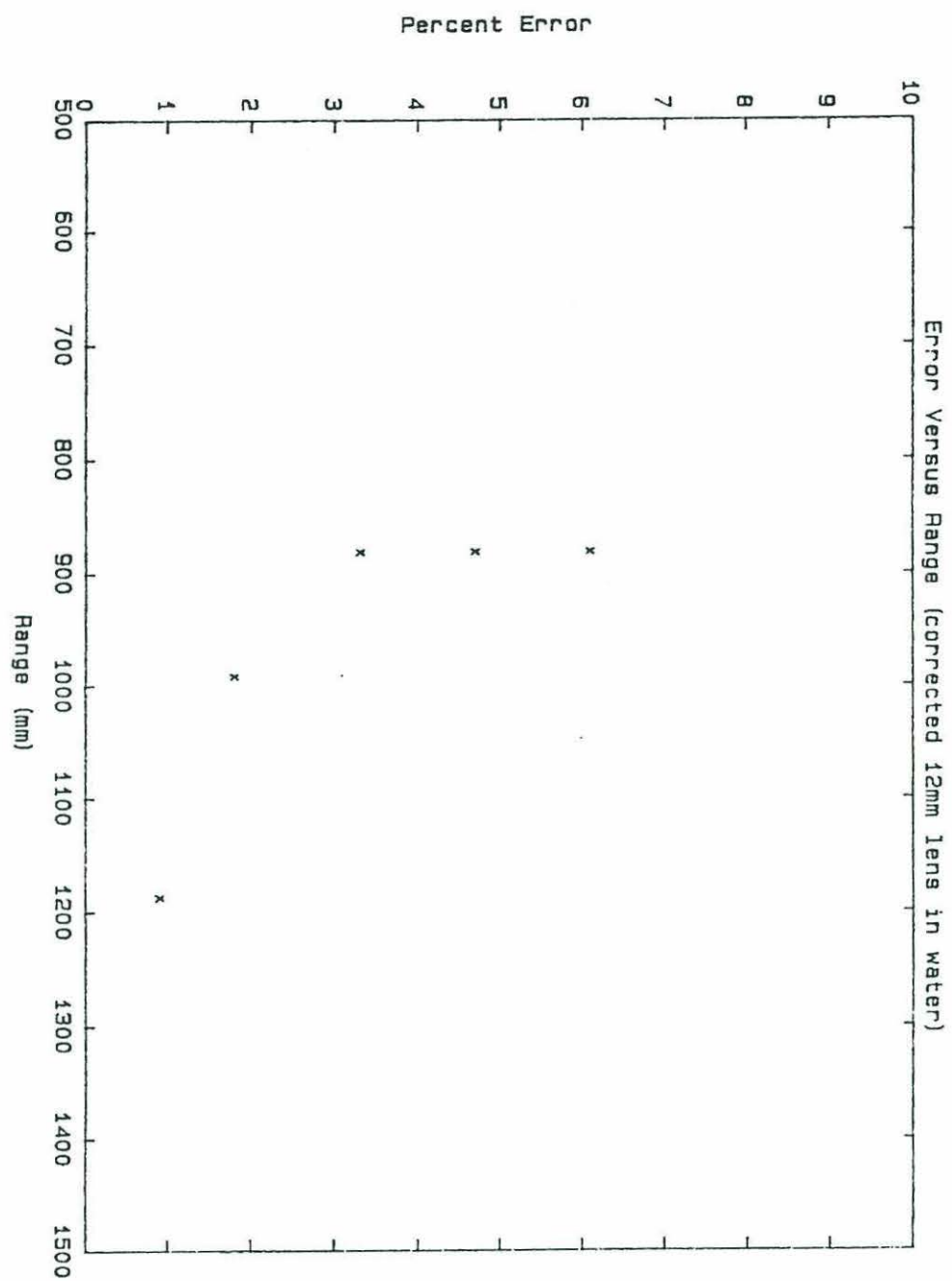


Figure 15

Appendix A 4.8mm lens in air

$$k_1 = 312.43537$$

$$k_2 = 184.29213$$

<u>x_i</u> (mm)	<u>y_i</u> (mm)
-75.0	150.0
0.0	212.132
106.066	75.0
75.0	0.0

Test 1: camera at (0,0,-530mm)

<u>u_i</u> (pixel)	<u>v_i</u> (pixel)
192	97
236	76
298	123
281	149

Result: $X_s = -55.9$, $Y_s = 48.5$, $Z_s = -525.2$

Test 2: camera at (-265,0,-459)

u_i	v_i
194	93
235	76
287	125
272	149

Result: $X_s = -259.2$, $Y_s = -50.1$, $Z_s = -470.0$

Test 3: camera at (-375,0,-375)

u_i	v_i
200	91
235	75
276	126
265	149

Result: $X_s = -370.7$, $Y_s = -28.5$, $Z_s = -389.8$

Test 4: camera at (-459,0,-265)

u_i	v_i
210	90
235	76
263	126
256	149

Result: $X_s = -455.9$, $Y_s = 0.9$, $Z_s = -264.6$

Test 5: camera at (0,0,-1030)

u_i	v_i
209	123
232	112
264	136
255	150

Result: $X_s = -22.9$, $Y_s = 25.6$, $Z_s = -1051.6$

Test 6: camera at (-515,0,-892)

u_i	v_i
211	122
231	112
259	137
252	150

Result: $X_s = -516.5$, $Y_s = -41.0$, $Z_s = -874.8$

Test 7: camera at (0,255,-530)

u_i	v_i
212	114
255	93
313	139
292	161

Result: $X_s = 4.5$, $Y_s = 256.4$, $Z_s = -559.8$

Test 8: camera at (-265,255,-459)

u_i	v_i
215	117
255	93
302	134
286	156

Result: $X_s = -256.5$, $Y_s = 255.5$, $Z_s = -511.4$

Test 9: camera at (-375,255,-375)

u_i	v_i
220	119
254	93
290	132
279	154

Result: $X_s = -385.1$, $Y_s = 269.7$, $Z_s = -382.4$

Test 10: camera at (0,456,-530)

u_i	v_i
198	144
236	128
283	162
266	178

Result: $X_s = -34.5$, $Y_s = 480.3$, $Z_s = -545.8$

Test 11: camera at (-265,456,-459)

u_i	v_i
200	149
235	128
275	154
260	171

Result: $X_s = -270.3$, $Y_s = 503.7$, $Z_s = -484.3$

Appendix B 8mm lens in air

$$k_1 = 545.59157$$

$$k_2 = 323.03207$$

Test 1: camera at (-220,82,-381)

u_i	v_i
165	94
253	51
362	152
335	205

Result: $X_s = -214.2$, $Y_s = 81.3$, $Z_s = -374.7$

Test 2: camera at (-311,82,-311)

u_i	v_i
174	91
248	51
332	152
314	205

Result: $X_s = -307.2$, $Y_s = 83.6$, $Z_s = -291.6$

Test 3: camera at (-420,70,-727)

u_i	v_i
220	121
262	98
319	150
305	178

Result: $X_s = -453.8$, $Y_s = 71.4$, $Z_s = -703.4$

Test 4: camera at (-594,105,-594)

u_i	v_i
226	120
262	98
309	149
297	177

Result: $X_s = -606.1$, $Y_s = 109.7$, $Z_s = -634.8$

Appendix C 12mm lens in air

$$k_1 = 818.38735$$

$$k_2 = 484.76785$$

Test 1: camera at (93,45,-540)

u_i	v_i
147	97
254	41
415	159
369	227

Result: $X_s = 95.5$, $Y_s = 44.9$, $Z_s = -541.3$

Test 2: camera at (160,110,-590)

u_i	v_i
161	100
259	49
405	158
364	220

Result: $X_s = 159.2$, $Y_s = 111.0$, $Z_s = -591.2$

Test 3: camera at (160,90,-690)

u_i	v_i
172	108
256	64
382	157
347	211

Result: $X_s = 166.4$, $Y_s = 89.6$, $Z_s = -686.1$

Test 4: camera at (-280,70,-485)

u_i	v_i
148	92
247	41
379	159
346	222

Result: $X_s = -288.9$, $Y_s = 68.5$, $Z_s = -483.8$

Test 5: camera at (-396,70,-396)

u_i	v_i
164	90
247	41
350	159
327	220

Result: $X_s = -396.0$, $Y_s = 77.0$, $Z_s = -385.9$

Appendix D 8mm lens in water

Test 1: camera at (-290,130,-502)

u_i	v_i
168	105
260	63
362	168
326	223

Result: $x_s = -286.9$, $y_s = 134.9$, $z_s = -503.8$

Test 2: camera at (-410,115,-410)

u_i	v_i
177	104
254	63
328	166
301	219

Result: $x_s = -409.6$, $y_s = 111.5$, $z_s = -393.6$

Test 3: camera at (200,105,-965)

u_i	v_i
215	123
270	97
342	160
315	192

Result: $X_s = 244.3$, $Y_s = 124.3$, $Z_s = -966.2$

Appendix E 12mm lens in water

Test 1: camera at (200,105,-965)

u_i	v_i
163	87
248	48
358	143
318	192

Result: $X_s = 187.6$, $Y_s = 95.9$, $Z_s = -955.2$

Test 2: camera at (200,105,-1165)

u_i	v_i
182	99
253	67
343	146
309	186

Result: $X_s = 200.5$, $Y_s = 115.0$, $Z_s = -1170.0$

Test 3: camera at (-440,30,-762)

u_i	v_i
163	76
250	36
351	142
315	194

Result: $X_s = -473.1$, $Y_s = 5.1$, $Z_s = -764.3$

Test 4: camera at (-622,30,-622)

u_i	v_i
187	75
260	36
338	142
309	194

Result: $X_s = -643.6$, $Y_s = 43.8$, $Z_s = -607.6$

Test 5: camera at (-762,30,-440)

u_i	v_i
218	75
267	36
314	142
295	193

Result: $X_s = -787.6$, $Y_s = 38.5$, $Z_s = -393.9$

Bibliography

- Duda, R. O., P. E. Hart, Pattern Recognition and Scene Analysis, New York, Wiley, 1973.
- Fischler, M. A., R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Proceedings Image Understanding Workshop, April 1980, pp 71-88.
- Ganapathy, S., "Decomposition of Transformation Matrices for Robot Vision", Proceedings of the First International Conference on Robotics, March 1984, pp 130-139.
- George, J. D., G. I. Lythgoe, J. N. Lythgoe, Underwater Photography and Television for Scientists, Oxford, Clarendon Press, 1985.
- Haralick, R. M., "Using Perspective Transformations in Scene Analysis", Computer Graphics and Image Processing, 13(1980), pp 191-221.
- Horaud, R., B. Conio, O. Leboulleux, B. Lacolle, "An Analytic Solution for the Perspective 4-Point Problem", Computer Vision, Graphics, and Image Processing, 47(1989), No 1, pp 33-44.
- Huang, T. S., R. Y. Tsai, "Image Sequence Analysis: Motion Estimation", Image Sequence Analysis, New York, Springer-Verlag, 1981, pp 1-18.
- Kanattani, K., "Camera Rotation Invariance", Computer Vision, Graphics, and Image Processing, 39(1987), No 3, pp 328-354.
- Lowe, D. G., "Solving for the Parameters of Object Models From Image Descriptions", Proceedings Image Understanding Workshop, April 1980, pp 121-127.
- McNeil, G. T., "Underwater Photography", Proceedings of the Society of Photo-optical Instrumentation Engineers, 24(1971), pp 117-131.
- Nasr, H., B. Bhanu, "Landmark Recognition for Autonomous Mobile Robots", Proceedings IEEE International Conference on Robotics and Automation, 1988, pp 1218-1223.

Roach, J. E., J. K. Aggarwal, "Determining the Movement of Objects from a Sequence of Images", IEEE Transactions on PAMI, November 1976, pp 554-562.

Sugihara, K., "Some Location Problems for Robot Navigation Using a Single Camera", Computer Vision, Graphics, and Image Processing, 42(1988), No 1, pp 112-129.

Sutherland, I. E., "Three-Dimensional Data Input by Tablet", Proceedings IEEE, 62(1974), No 4, pp 453-461.